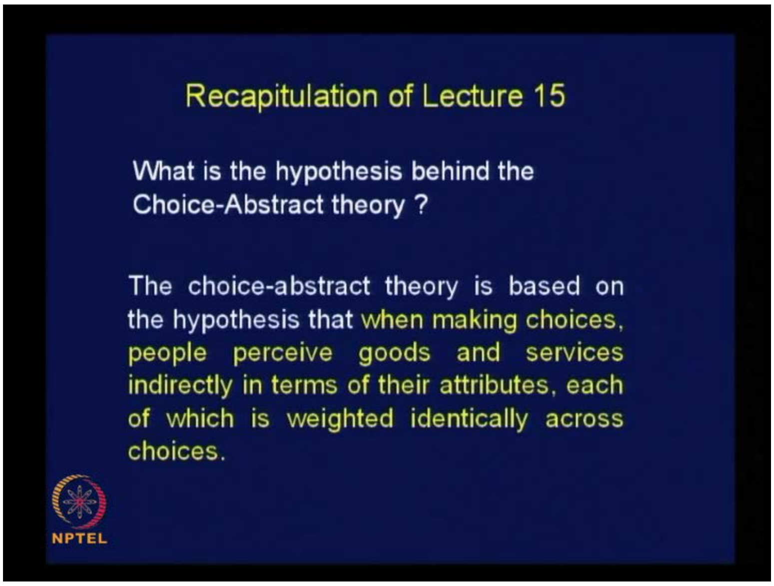
Urban Transportation Planning
Prof. Dr. V. Thamizh Arasan
Department of Civil Engineering
Indian Institute of Technology, Madras

Lecture No. #16
Modal Split Analysis Contd.

This is a lecture sixteen on urban transportation planning. We will continue our discussion on modal split analysis in this class, let us recall what we did in the previous class briefly, so that, there is continuity of the discussion. You may recall, we mainly discussed about attribute specific utility function in mode choice analysis. The reason or the advantage of attribute specific mode choice functions is that, it would be possible for us to introduce any mode into your utility function process. You will have a single utility function, and all model characteristics can be incorporated easily; that is the advantage of attribute specific utility functions. In that context, we discussed about some problems related to utility function mainly with regard to the constant term, how to account for the value of the constant term, when we deal with a set of alternative modes. Then, we discussed about incorporating the effect of trips specific advantages for a particular mode in the utility function.

We also discussed about the possibility of incorporating, the socioeconomic characteristics of travelers in the utility function. And finally, we discussed about the randomness of the utility itself, because individuals perceive utility of modes in their own way; that is how utility of a particular mode for different individuals happens to be different. We cannot fix a constant value as utility for a particular mode, it varies from individual to individual. Hence, it is appropriate to bringing this variation by introducing some randomness in the utility function. In that process, we discussed about the introduction of a random term with the symbol epsilon, and finally we discussed about the possible distribution for epsilon.

(Refer Slide Time: 03:13)



Before, we proceed further in our discussion; let us try to check our self whether we have captured the essence of the previous lecture, by posing a set of questions, I hope some of would be able to answer these questions. The first question is this, what is the hypothesis behind the choice abstract theory hypothesis behind the choice abstract theory, any suggestion

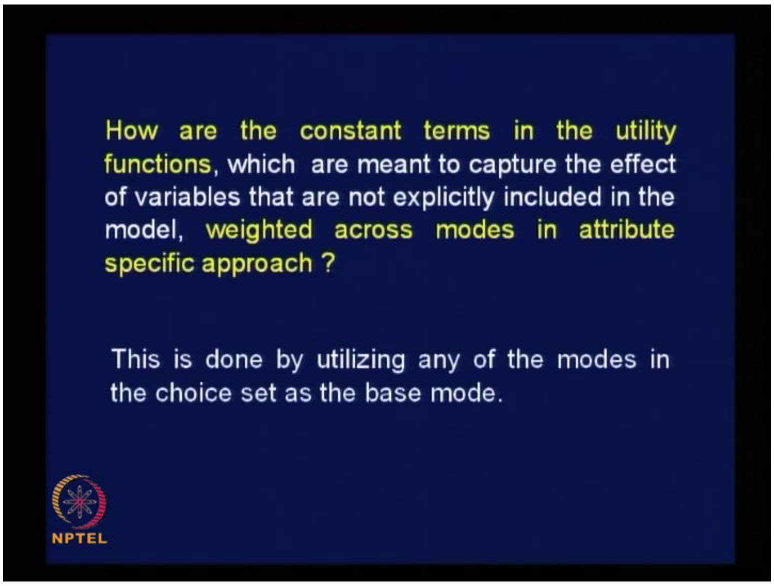
The choice of a particular mode depends upon the attributes of it, attributes of a particular mode.

Yes.

Based on that people will select it.

Fine, that is the the result or effect of choice abstract theory, the hypothesis can be explained in an abstract form as follows. The choice abstract theory is based on the hypothesis that when making choices, people perceive goods and services indirectly in terms of their attributes, each of which is waited identically across choices. In that context only we discussed about, choice of cell phone services based on a set of attributes choice of car. For example, if the engine size is 1.2 liters for Tata make or Hyundai make or Maruthi Suzuki, we do not differentiate between makes, when the engine size is same, we perceive that utility of the car of these engine size irrespective of the make is same that is the meaning implied here, we depend mainly on attributes not specific to a particular good or service. So this is the hypothesis behind the choice abstract theory.

(Refer Slide Time: 05:10)

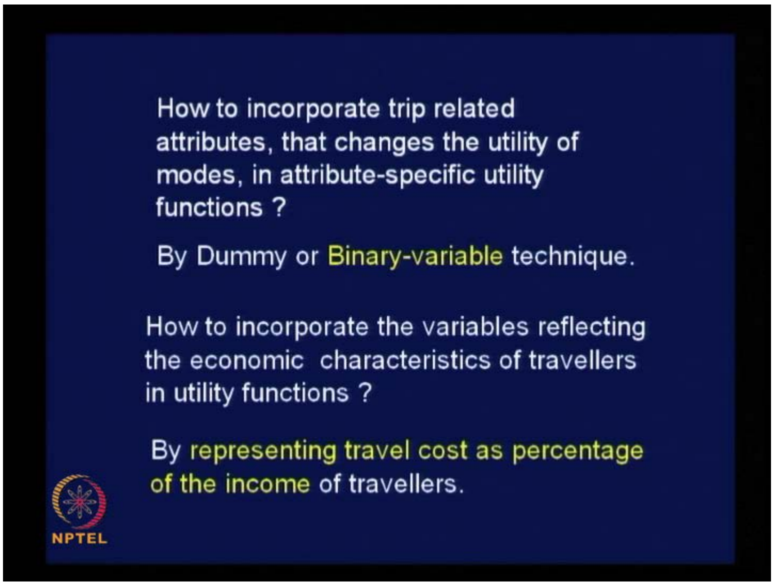


Next question is this, how are the constant terms in the utility functions, which are meant to capture the effect of variables that are not explicitly included in the modal, weighted across modes in attribute specific approach. How are the constant terms in the utility functions, weighted across modes in attribute specific approach, why should we wait the constant terms across modes because we have only one general function incorporating all the attributes into it and only difference we want to make is based on the constant term, which stands for the unexplained part of the dependent variable right and that, if not change, will have the same effect for all the alternatives which may not be realistic. So the unexplained part will be different for different alternatives, we want to bring in that difference, how do we do that?

Taking one as this function and explaining in terms of this.

Good, taking one mode is a base mode, we make the constant term to be 0 for the base mode and then we fix the relative values for the other modes, that is how, we manage this particular situation then it facilitates variation of utility of different modes, even though the complete variation is not explained by the set of attributes considered for utility function right. So this is done by utilizing any of the modes, in the choice set as the base mode. Once you do this, you will be able to differentiate, the extent to which the unexplained part is considered in the utility functions for the different modes.

(Refer Slide Time: 07:28)



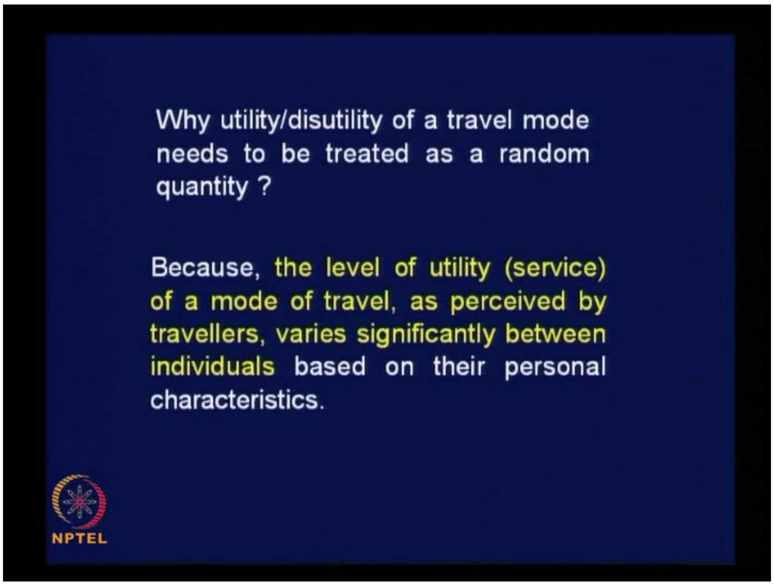
The next question is this, how to incorporate trip related attributes, that changes the utility of modes, in attribute-specific utility functions. First we need to understand that tris trips specific attributes may change the value of utility of a particular mode, do you agree with this statement. Unless we agree on this, we will not be able to answer this question, can specific trips change the utility of a given mode yes or no, I think most of you respond by say yes. So trip or the nature of trip can change the level of service or utility of a particular mode.

In this context only, you may recall, that we discussed about trips towards CBD, trips connected to CBD and trips not connected to CBD. In certain specific situation trips connected to CBD, may have added utility for transit service because of restriction in parking (( )) CBD areas the personal modes become less attractive, utility gets reduced. Now the question is, how do we account for this change, utility of a particular mode, when a trip are of different nature, what is the technique adopted for this purpose. You may recall, we discussed about binary variable technique right. So when a trip is connected to CBD, your sign value one for that particular variable otherwise 0. So that is, how we incorporate this aspect also in the attribute-specific utility functions. By dummy or Binary-variable technique, that is the answer for this question is that clear.

The next question is this, how to incorporate the variable reflecting the economic characteristics of travelers in utility function, we would like to incorporate the effect of economic status of travelers in the utility function. It is essential because the choice of mode by travelers, it depends on their economic status, how do we do this, any suggestion, you may recall that we discussed about, expressing travel cost as a percentage of the income of the traveler, what is the purpose. Once you express the travel cost as a percentage of income, for a higher income group, the percentage is going to be very small; for middle income group, percentage will be having some medium value; low income group, the percentages going to be very high, that reflects the pinch of (( )) experienced by the different groups of travelers based on their economic status is not it.

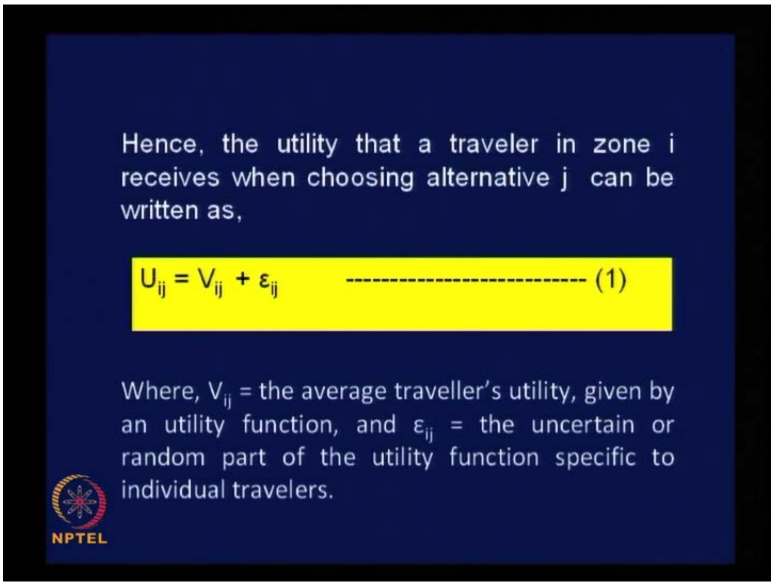
So that is a very important factor affecting more choice, so this has to been cooperated, by this particular methodology, so, this is the answer for this question. By representing travel cost as percentage of the income of travelers, do not give absolute values represent it, as a percentage of income. This implies, that you are going to categorize travelers in to different income groups, while developing utility functions, that is the point mentioned. In that context, in the previous class I also mentioned about the possibility of introducing other social and demographic factors in the utility function, like the effect of age in mode choice, what is the technique adopted for introducing effect of age, in the utility function, how do we do that? Use binary variable technique, fix an age limit. People below that age will have some value in the binary variable and the above that age limit will have another value is not it or if you want have three categories it is possible, you have three different numbers, for the three categories of people based on age right. So that is how, we should think of introducing all possible factors that might influence mode choice.

(Refer Slide Time: 13:08)



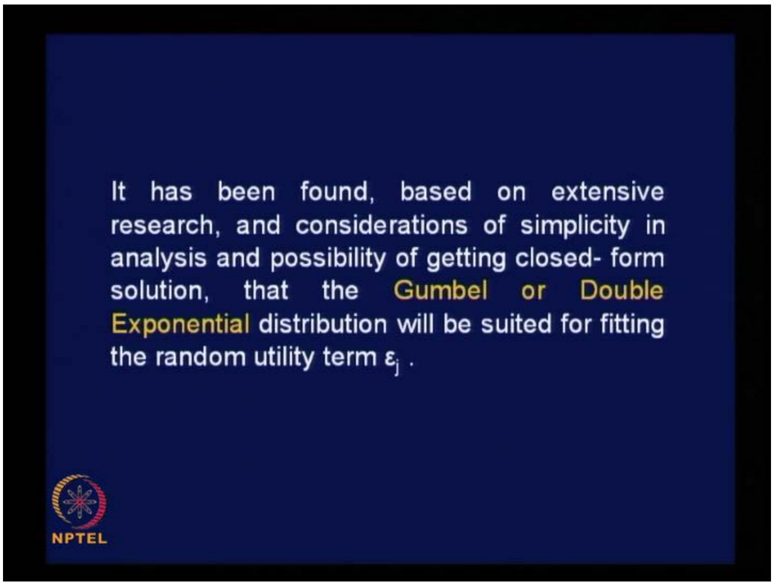
Why utility or disutility of a travel mode needs to be treated as a random quantity, since we have discussed about this, just now I thought will straight away look at the answer. The reason is this, because the level of utility or the service of a mode of travel as perceived by travelers varies significantly between individuals right, based on their personal characteristics, that is why there is a need to treat utility as a random quantity, and not a fixed quantity. On this basis only we proceeded to formulate a function for utility incorporating this randomness, is not it.

(Refer Slide Time: 14:09)



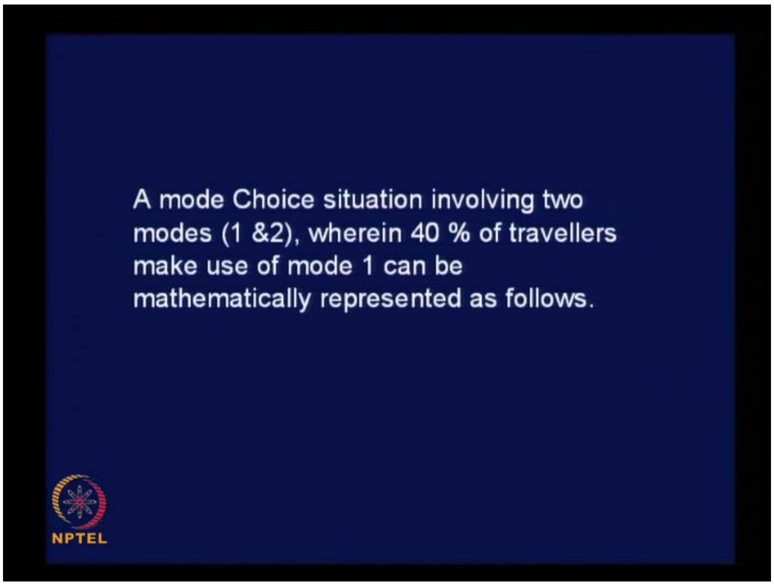
Proceeding further, we can say now, the utility that a traveler in zone i receives when choosing alternative j can be written as follows, as we have seen in the previous class U i j is equal to V i j plus epsilon i j, where V i j is the average travelers utility assumed to be constant for a given mode, given by an utility function and epsilon i j is the uncertain or random part of the utility function specific to individual travelers.

(Refer Slide Time: 15:05)



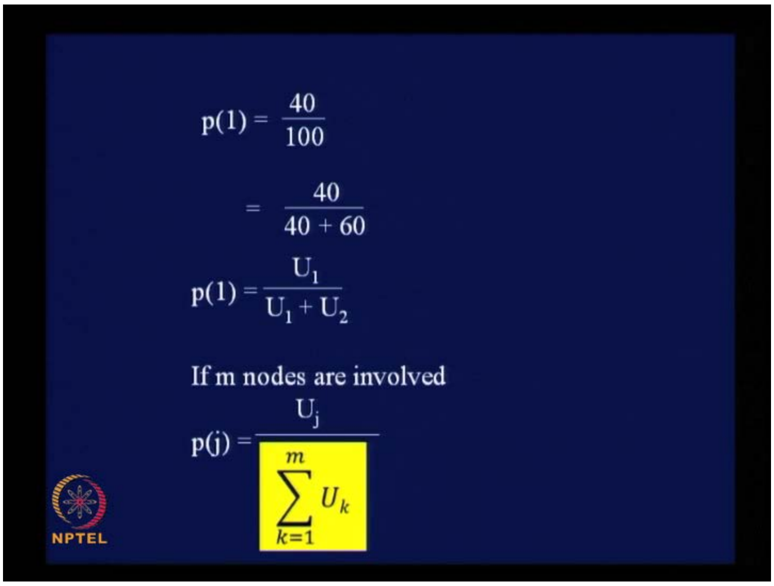
And then, we were discussing about the possible functional form for epsilon i j. So that, we are able to quantify U i j ultimately to analyze a mode choice and in the process of discussion we found that, several researches have recommended Gumbel or Double exponential distribution as more appropriate distribution for epsilon i j this is based on consideration of simplicity in analysis and possibility of getting closed-form solution. So based on this assumption only we found, that it is possible to treat utility as a random quantity and still quantify utility for various alternative modes.

(Refer Slide Time: 16:01)



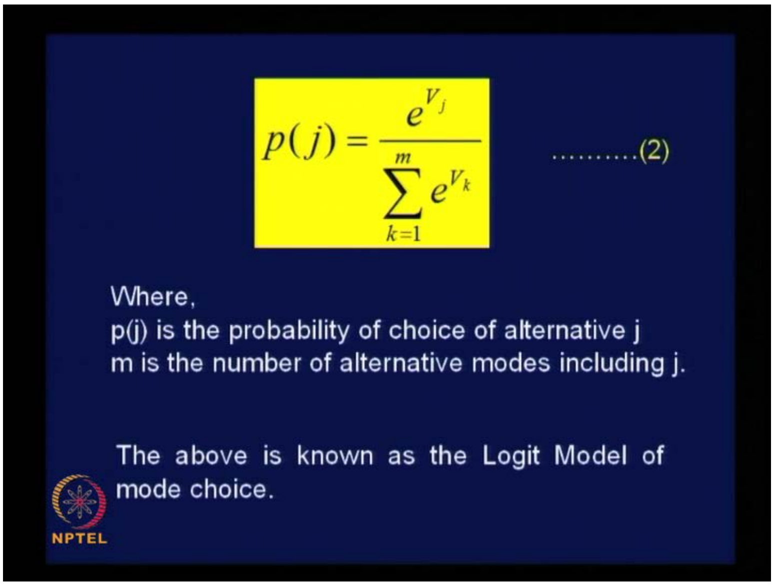
To understand the concept little more clearly, let us consider a very simple situation involving mode choice between two modes, mode 1 and 2, where in 40 percent of the travelers make use of mode 1 and how to express this mathematically, we have a situation, where there are two alternative modes available and 40 percent of travelers make use of mode 1, can you express this situation mathematically.

(Refer Slide Time: 16:52)



This is how we can mathematically express the situation, probability of choice of 1 is 40 by 100, 40 percent of people make use of mode 1, so that will be equal to 0.4 probability of choice of mode 1, here it is 0.4 and we can also write the same thing as 40 divided by 40 plus 60 implying the share of two modes that is what is taken the denominator is not it 100 implies 40 plus 60. So in general, if we take the 40 percent as utility of mode 1, you can say U 1 to be equal to or probability of choice of mode 1 is equal to U 1 divided by U1 plus U 2 is not it. So, if there are m alternatives, we are going to write probability of choice of j. For example in general, will be equal to U j divided by sigma k is equal to 1 to m U k is not it, some of utility of all the modes agreed and when you assume gumbel distribution for epsilon i j and try to derive an equation for probability of choice of a particular mode, you will get a similar formulation, but little differently.

(Refer Slide Time: 18:45)

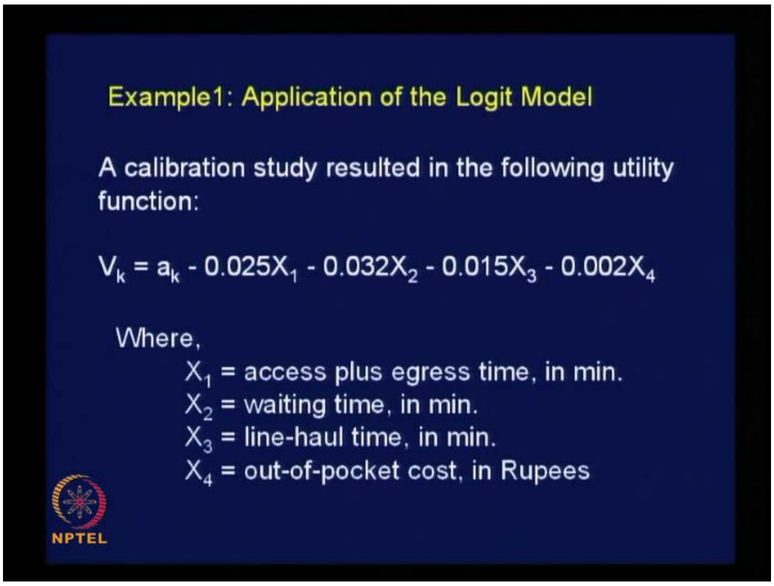


I will directly give you the result of the equation or the function, probability of choice of j is given as e power V j divided by sigma k is equal to 1 to m e power V k, what is the difference between the previous equation and this equation. Instead of U, we are introducing e power V both in the numerator and the denominator right. So, based on the probabilities aspect of probability state theory, if you derive a formula, in general for probability of choice of any mode or a given mode based on the assumption that epsilon i j follows gumbel distribution or double exponential distribution, we end up with these result.

The derivation of a formula is beyond the scope of this particular course, so that is why, I am not bringing in the derivation part, which needs quite a good amount of background, knowledge in probability and statistics. Those who are interested, can refer any one good book dealing with probability of choice of alternatives, you will be able to get information about the derivation of this equation right what I have given here is just the result, but you should appreciate that, it is similar to what we have seen, the only difference is when you derive, we find that it is possible to incorporate the randomness, without bringing in epsilon i j in the equation, what do use is the fixed part of the utility V j and V k that is a interesting aspect of it. If you go through the derivation process, you will understand how epsilon i j gets eliminated, when we transform into this particular functional form right, where p j is a probability of choice of alternative j and m is the number of alternative modes including j. So suggestion made here is, we have to take this equation for granted right and this equation is also named as logit model of mode choice, why it is named as logit model of mode choice.

You may recall the shape of the curve that we assumed for distribution of epsilon i j, that curve is also named as logistic curve, another name given by researchers is logistic curve, since logistic curve is taken as a basis for the derivation. The model itself is named as logit model of mode choice. So, logit model of mode choice is nothing but, this equation P j is or p f j is equal to e power V j divided by sigma k is equal to 1 to m e power V k. The above is known as the logit model of mode choice. Now, let us just see how to apply this logit model for a given situation, we will directly take up a small numerical example and see how logit model can be applied to analyze mode choice clear.

(Refer Slide Time: 23:04)



Let us consider this example, a calibration study resulted in the following utility function, V k in general is found to be equal to a k minus 0.025 X 1 minus 0.032 X 2 minus 0.015 X 3 minus 0.002 X 4, where X 1 is equal to access plus egress time in minutes, I think you know clearly now what we really mean by access time and egress time, time to the transit station is access time, time from the transit station to the destination is egress time; X 2 waiting time in minutes; X 3 line haul time in minutes, line haul time or in vehicle time both the same and X 4 out of pocket cost in rupees about which we had enough discussion about what we really mean by out of cot cost expenditure in respect of transport. So, please note that all these variables are related to mostly time and the last one is related to cost. So here utility of a given mode has been decided based on a set of time components and the cost only two attributes mainly, but the first attribute is categorized into three types, why categorization into three types why not take the total time.

Because each and every mode has different time

Each and every

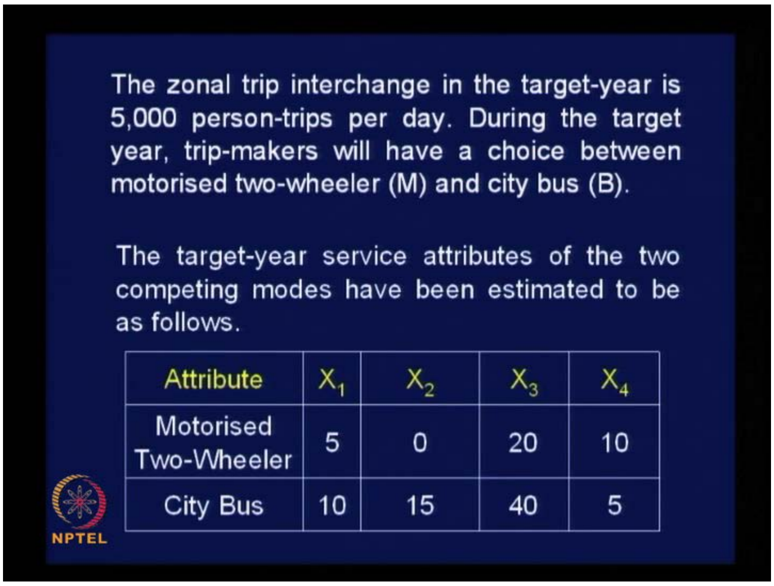
Every mode has a different time

Ok

So, it has to be classified, so that mode can be easily analysed.

That is a point, because when we consider different alternative modes these components are going to be different. In certain cases, you may find some components may be even 0. For example, if you take personal mode there is no waiting time, if we take transit there is waiting time component, that is why we need to split the journey time into different component. So that, the real utility of different modes are brought into your utility function right, so this is a calibrated utility function, why negative sign for all the variables, because the author as treated these utilities as really disutility's they spend time, they spend money why not term it as disutility. In that context, you get negative sign for all this variables right. So this is the calibrated utility function.

(Refer Slide Time: 26:32)

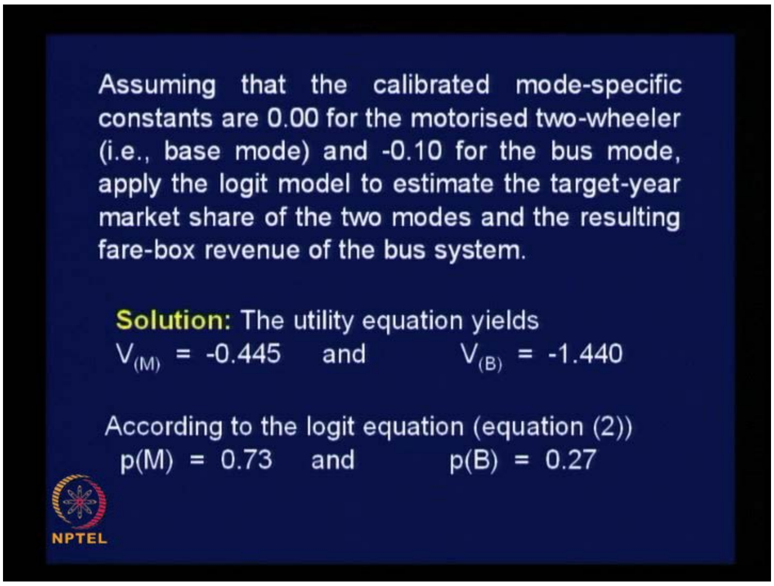


And the zonal trip interchange in the target target-year is 5,000 person-trips per day, please note, this is zonal trip interchange, trip interchange between a pair of zones. We do not consider entire urban area; it is specific to a particular travel market. During the target year, trip-makers will have a choice between motorised two-wheelers and city bus. This targets a particular socioeconomic group, where they have specific choice of only two alternatives.

The target-year service attributes of the two competing modes have been estimated to be as follows, target year or horizon year, these are the values of X 1, X 2, X 3 and X 4 for the two modes. For motorised two-wheeler X 1 is what, as per our initial understanding of the variables, X 1 is axis and egress time, 5 the unit is same in both the cases you can take it as five minutes; X 2 is waiting time; X 3 is line on time and X 4 is actual travel cost in rupees. So; obviously waiting time is 0 for mode cycle and line haul time is 20 minutes and cost let us say rupees 10.

It is a zonal pair travel, you can consider this as a cost for one trip as indicated here, for city bus, X 1 access egress time is more than X 2 there is waiting time 15 minutes and cost no, the third one is line haul time is 40 minutes and cost is less about half of the cost implication with motor cycle right. And we have to find out what, we have to get utility function or the value of utility for the two modes and then evaluate or estimate the probability of choice of motor cycle as well as bus that is our ultimate objective.

(Refer Slide Time: 29:38)

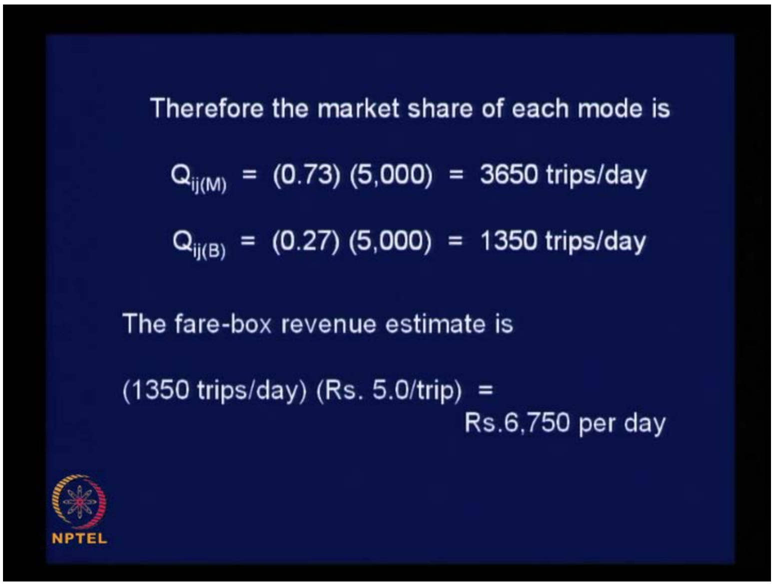


And, of course we can calculate few other things also that is what is as here, assuming that the calibrated mode-specific constants are 0.00 for motorized two-wheelers, which is taken as a base mode, that is why 0 and minus 0.10 for bus mode, apply the logit model to estimate the target-year market share of the two modes, market share is asked for, how many travelers out of the total will use motor cycle and how many will use bus as to be estimated finally, and the resulting fare-box revenue of the bus system. Let us see last question to be answer, what will be the collection for the bus service, fare box revenue. Let us see how to go about doing this exercise using the utility function we have got the results here, the results are shown here, utility function is given to us earlier right.

Substitute the values of X 1, X 2, X 3 and X 4 corresponding to motor cycle right then you get the utility shown here as V M to be minus 0.445 or disutility related to motor cycle is minus 0.445 and using the function, if you calculate disutility for bus, it is minus 1.440 is that clear. We have the equation given earlier, substitute the values of X 1, X 2, X 3 and X 4 corresponding to the two modes and you get the values of utility for the two modes. Are you able to appreciate this point, just I will go back to the equation, this is a equation given, calibrated utility function. And if you substitute a k to be 0 right, for motorcycle and substitute the values of X 1, X 2, X 3 and X 4 corresponding to motorcycle, which is 5, 0, 20 and 10 will get the value of V k is not it. So that is how we got values of V k named here as V M for motorcycle as minus 0.445 and for bus named here as V B as minus 1.440 utilities are calculated now, of course this is fixed part of the utility and then, substitute this values in the logit model is not it.

You can get probability of choice of motorcycle as 0.73 and probability of choice of bus as 0.27 obviously, are you able to appreciate this point, how do you get this probabilities. We have the values of V M and V B is not it and use the logit model which you have seen here, probability of choice of motorcycle p M p M here, is going to be e power V M divided by e power V M e power V B, the values of V M and V B are known. So, we can get the value of p M as well as p B, if you want estimate p B, p B is equal to e power V B divided by e power V B plus e power V M is not it, that is what we do to get this result, that is how we get probability of choice of motorcycle as 0.73 and probability of choice of bus as obviously 0.27. We know the total number of trips, knowing the probability you can calculate the actual model share is not it, by multiplying the total number by the corresponding probabilities.

(Refer Slide Time: 34:15)

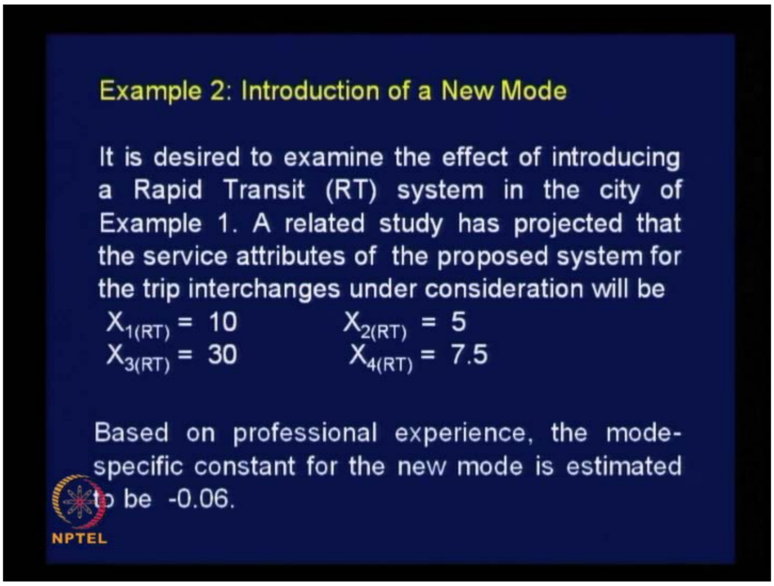


So the market share of each mode is then Q i j M because it is origin destination specific, so we can say Q i j by motorcycle M is equal to probability value is 0.73 and total number of trip interchanges on 5,000, so actual model share is 3650 trips per day made by motorcycle between this particular set of origin and destination and Q i j B by bus 0.27 into 5,000 that is equal to 1350 trips per day right, from this information can we calculate the fare-box revenue, total number of travelers is known, that is that is what is meant here, as Q i j B model share is known to us so many trips, these trips are person trips not vehicular trips, person trips we are dealing with person trips, what is the cost of travel by bus.

(( ))

Out of pocket costs that was indicated as X 4 in the data, which is ten for motorcycle and five for bus, so this number into five will give you the fare-box revenue. The fare-box revenue estimate is 1350 trips per day into rupees 5 per trip, that is equal to 6,750 rupees per day for the bus operator. So it is a very simple example, but at the same time gives you a revealing application in practice. So this is what we do, by way of applying logit model of mode choice.

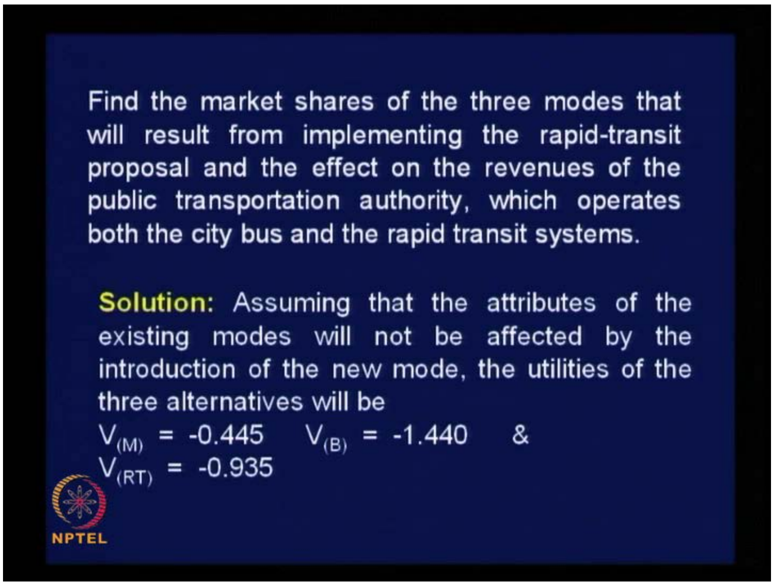
(Refer Slide Time: 36:35)



We will continue the extension of this example, as example 2, introduction of a new mode, let us say we are working the mode choice in a rear for a horizon air condition right. Now it is desired to examine the effect of introducing a rapid transit right, in the city of example 1 in the previous case, where we consider two modes, a related study has projected that the service attributes of the proposed system, rapid transit system for the trip interchanges under consideration between the same zonal pair will be as follows as given here X 1 for rapid transit is going to be 10, what is X 1 access and egress time is not it, is going to be 10 minutes; X 2 is waiting time is going to be on the average 5 minutes; X 3 line haul time is going to be 30 minutes, you may recall line haul time for bus was 40 minutes, whereas here it is slightly less and it is estimated to cost about 7.5 rupees per trip slightly more than bus, but less than motor cycle is not it. So, these are the parameter values or values corresponding to various attributes that we are brought into our utility function.

Now, we have how many alternatives, after introduction of rapid transit, three alternative modes, there will be three utility modes and we need to calculate three probabilities and then estimate the model share right. And this is another important information given, based on professional experience, the mode specific constant for the new mode is estimated to be minus 0.06. This is also required because we need to accurately estimate utility for this mode; this minus 0.06 is based on the same assumption, that motorcycle is treated as a base mode right. And there is no change in the constant term consider for bus, without making any change in those constant terms, it is found by experience that the constant term for this mode can be taken as minus 0.06.

(Refer Slide Time: 39:47)



Now this is the question, find the market share of the three modes that will result from implementing the rapid-transit proposal and the effect on the revenue of the public transportation authority, which operates both the city bus and the rapid transit system, which is somewhat peculiar to our condition, here rail and bus services are operated as separate entities, there is not even any coordination between these two modes, that is the way we operate our transit services, which is ultimately becoming unfriendly for the users, it is not really user friendly.

If these two modes as indicated here are operated by a single public transit authority, there will be several advantages managing the system will be much easier in the operators point of view and in the users point of view, users can treat both this modes as a single mode, as for as travel is concern, users can be given a single ticket and they can make seamless travel an waiting time can be reduced, because the same operator is operating both rapid transit as well as bus service, they can schedule the bus services in such a way that, when the bus arrives the waiting time is just sufficient to go and take the rapid transit service train service.

All these things are possible that is, what is happening in most cities of developed countries. Single ticketing systems and all transit services under single authority, even though practically operators could be separate entities, as far as revenue collection is concerned. It is done by a single authority and then shared by the operators based on the actual demand, for each of these modes as well as level of utilization, maintenance commitments and so on. That is how this transit system is being operated in developed countries; hopefully we get such a situation shortly in our country. People are talking about, at least conceptually with regard to integration of all the modes and developing a single authority for managing transit system and so on right, getting back to the question.

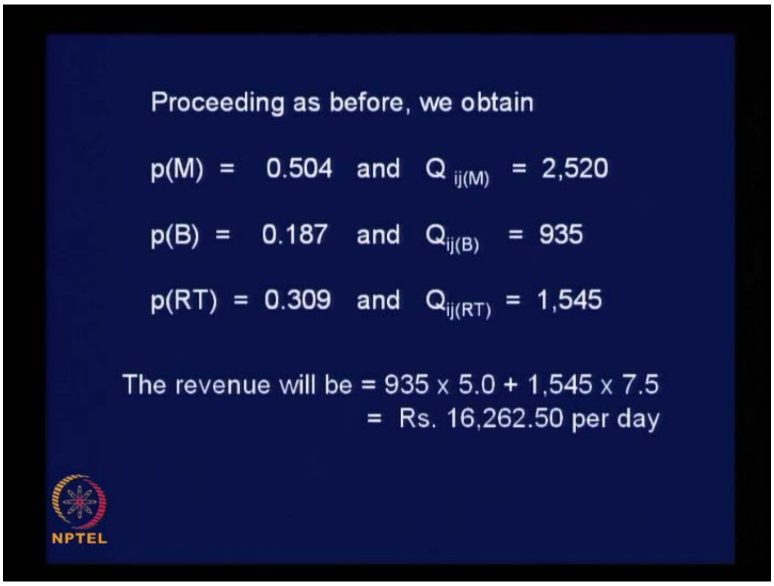
Let us get the solution as follows, assuming that the attributes of existing modes will not be affected by the introduction of the new mode, why this assumption, assuming that the attributes of the existing modes will not be affected by the introduction of the new mode, what are the attributes of the existing modes, which are of our concern. The attributes, we are considered X 1, X 2, X 3 and X 4 and these values are assumed to remain unchanged, even after the introduction of a new mode, which need not necessarily be true in all the cases, which are the attributes of the exciting modes, that are likely to changed because of introduction of a new mode, any response

(( ))

Is there be a change, for example in X 1 of the modes the other two modes when you introduce a new mode, X 1 is axis and aggress time, will there be a change, when introduce a new mode may not be any change; X 2 is waiting time could there be change, when you introduce a new mode, waiting time of the previous the already existing modes are the likely to is there is there a possibility of change in the waiting time, for the other two modes may not be changes, unless the new mode is also introduced on some roads in parallel with existing bus service for example, then there will be a split in the demand and the bus operator might choose to operate or reduce the frequency of service, which will lead to increased waiting time is not it. So, there is a possibility of X 2 of existing modes changing, because of the introduction of a new mode about X 3 line haul time may not change much X 4 the cost.

If the introduction of the new mode creates competitiveness among the modes within the transit system then to compete for the share of the market, the bus operator might reduce the fair is not it. Competition is developed, so there is a likely hood of change in the value of X 4 for the existing modes right, but here we assume that all this attributes remain unaffected, for these case right, after the introduction of the new mode clear. And the based on this assumption, we can quantify utility value for the different modes, here we have three modes, so we estimate utility values as V M to be 0 minus 0.445 and V B minus 1.440 and V R T rapid transit minus 0.935. You can see the utility value stands between motorcycle and bus right then let us use the logit model of more choice and find the probability of choice of each of these modes.

(Refer Slide Time: 46:51)



I will be giving the result directly, proceeding as before in the previous example, we can get probability choice of motorcycle to be 0.504 and the actual market share will be 2,520 trips. And probability of choice of bus, will be 0.187 and the market share of bus will be 935. And probability of choice of rapid transit will be 0.309 and the market share will be 1,545 trips right. Now, we get a very clear picture of the share of the market by the three alternatives modes. Now to answer the final question, the revenue for the transit system operator, they need to multiply market share by the corresponding cost, bus share is 935 trips and cost is 5 rupees, so 935 into 5 plus market share of rapid transit 1,545 and the corresponding cost is 7.5 rupees, so total is 16,262 rupees 50 paisa per day, that will be the fare box revenue, for the transit system operator involving both bus and rapid transit system.

This is for a horizon years situation, that is very important, but we are using a calibrated utility function is not it, how are the utility function is calibrated, how do you calibrate utility function

(( ))

Using the base year data, you have to collect data initially to calibrate any function right, how do you collect the data and use the data for calibration of utility function, on the left hand side, you have utility of a particular mode is not it; on the right hand side, you have a constant term and then values of the X 1, X 2, X 3 etcetera representing the different attributes of alternative modes. In practice will be collecting data pertaining to all the independent variables X 1, X 2, X 3, X 4 and so on, also you will know the value of V 1, V 2, V 3 etcetera for base a condition is not it, how do you get the value of V 1?

(( ))

You will know the market share of different alternative modes right, let us say thirty percent of people are using mode one, forty percent mode two another thirty percent mode three see this model share reflects the utility, values of these modes is not it. So once you know this values, it is possible to calibrate any utility function, may be later on we will discuss in detail about calibration of such models. To summarize what we discussed in this class, we try to understand how we get the logit model of mode choice, from the principle that the utility function is a random quantity based on the assumption of randomness, we have seen that following systematic derivation procedure, it is possible to get the model that we just now seen as logit model of mode choice.

And we sort two numerical examples in the application of logit model of mode choice, the first example involved two alternative modes, and we understood clearly how to estimate the market share for two alternative modes. And of course, to calculate the fare box revenue for the transit system, and we also discussed about a situation, where a new mode is introduced, and the procedure of applying the logit model of mode choice, in estimating the market share for the different modes, including the possible newer mode that will be introduced in the horizon year. With this we were conclude our discussion for today, we will continue on this subject matter in the next class.